PARI: Policy-Driven Active Residual Intervention for Weakly Supervised Point Cloud Segmentation

Wenhao Gao[1*] **Sen Tao**[1*] **Jiawei Liu**[1†] **Jiwei Feng**[1] **Yongchao Xu**[1]
Junfeng Wang[1] **Tao Jiang**[2] **Jiangbo Ai**[2] **Jin Zhang**[2] **Zheng-Jun Zha**[1]

[1]MoE Key Lab of BIPC, University of Science and Technology of China, Hefei, China
[2]Alibaba Group, Hangzhou, China

{gaowh,taosen,jerryfeng,yongchaoxu,jfwang2002}@mail.ustc.edu.cn
{jwliu6,zhazj}@ustc.edu.cn {cigarette.jt,aijiangbo.ajb,zj146372}@alibaba-inc.com

Abstract

3D Weakly Supervised Semantic Segmentation (3D-WSSS) on point clouds aims to learn dense 3D semantics from extremely sparse point-level annotations. Existing methods follow a passive paradigm based on pseudo-label propagation or consistency regularization, leading to two key issues: 1) *Lack of targeted intervention*, where erroneous predictions are not explicitly identified and corrected, resulting in severe error accumulation; 2) *Limited contextual awareness*, where supervision is restricted to point- or view-level self-augmentation consistency, failing to capture the intricate contextual correlations embedded within the scene. To address these limitations, we propose a Policy-driven Active Residual Intervention (PARI) framework for 3D-WSSS, where an Active Intervention Agent (AIA) explicitly localizes high-risk points under the guidance of a Confidence-aware Reward Shaping (CRS) strategy, and a Bias-aware Refinement Module (BRM) subsequently executes targeted refinement by exploiting rich spatial and semantic scene contexts. Specifically, AIA actively localizes high-risk points from uncertainty maps and prediction priors, while BRM refines them through residual correction based on hybrid spatial-feature context. CRS introduces a high-confidence suppression reward that discourages unnecessary intervention on reliable predictions, and a low-confidence exploration reward that encourages intervention on uncertain predictions. Experiments on S3DIS and ScanNetV2 demonstrate strong performance across multiple weak supervision settings and consistent gains across the evaluated backbones. With only 1% point-level annotations, PARI achieves performance comparable to the reported fully supervised reference using the same backbone.

1 Introduction

Point cloud semantic segmentation is a fundamental task in 3D scene understanding, with broad applications in autonomous driving [14, 31, 55], robotic manipulation [4, 11, 29] and perception [48, 19, 3, 50]. In recent years, fully supervised deep learning methods have achieved remarkable progress on this task. However, these methods [24, 25, 39, 12] require dense point-wise annotations, while the acquisition and annotation of point clouds are extremely expensive, which severely limits their practical deployment. To alleviate this issue, 3D Weakly Supervised Semantic Segmentation (3D-WSSS) on point clouds has attracted increasing attention, aiming to learn dense 3D semantic representations from only a very small number of point-level annotations [37, 21].

*Equal contribution.
†Corresponding author.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

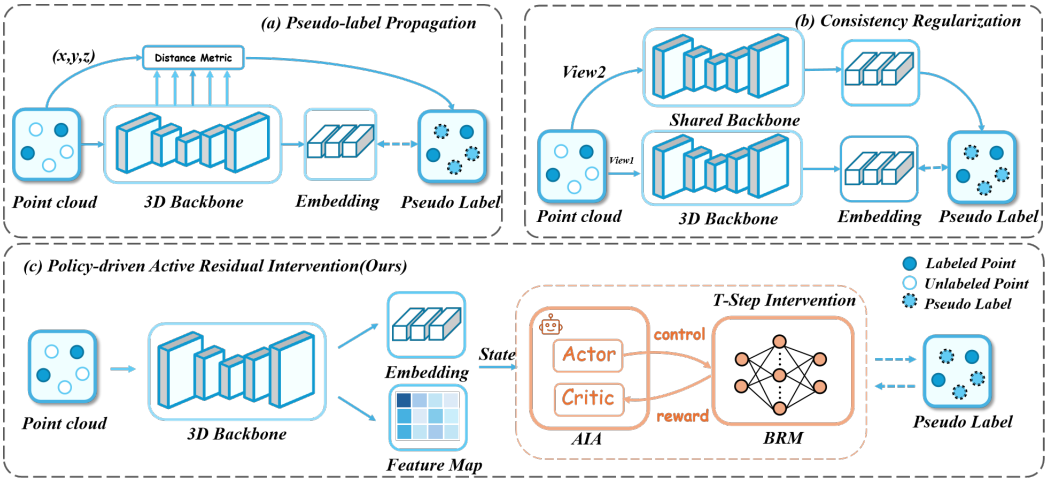

Figure 1: Comparison of weakly supervised point cloud segmentation paradigms. (a) Pseudo-label propagation passively densifies sparse supervision and amplifies errors. (b) Consistency regularization boosts robustness but lacks explicit intervention. (c) PARI performs policy-driven active residual intervention where AIA selects *where to intervene* and BRM determines *how to intervene*.

Existing methods alleviate label scarcity by densifying supervision signals. However, they largely follow a passive paradigm, which manifests in two main streams with distinct limitations: 1) Pseudo-label propagation methods (Fig. 1 (a)) propagate sparse annotations or high-confidence predictions to neighboring points. However, they suffer from **a lack of targeted intervention**, as erroneous predictions are not explicitly identified or corrected, causing incorrect semantics to be continuously propagated and resulting in severe error accumulation, particularly near semantic boundaries and in regions with complex geometries. 2) Consistency regularization approaches (Fig. 1 (b)) enforce global consistency across prediction distributions or feature representations between original and augmented views to improve robustness. However, they exhibit **limited contextual awareness**, as supervision is constrained to point- or view-level self-augmentation consistency, failing to capture the intricate contextual correlations within the scene.

To address these issues, we construct a Policy-driven Active Residual Intervention (PARI) framework for 3D-WSSS, where an Active Intervention Agent (AIA) and a Bias-aware Refinement Module (BRM) collaboratively identify and refine high-risk points in an iterative manner, while maintaining reliable predictions through a Confidence-aware Reward Shaping (CRS) strategy, as illustrated in Fig. 1 (c). Specifically, PARI adopts a two-stage training paradigm for 3D-WSSS. During Stage I, PARI trains a base segmentation network to produce initial semantic representations and predictions, serving as a stable yet imperfect prior for subsequent intervention. Building upon this prior, PARI formulates Stage II as two coupled questions: AIA addresses *where to intervene*, while BRM focuses on *how to intervene*. AIA generates an intervention mask based on the current environmental state (*i.e.*, the backbone features, uncertainty maps, and prediction priors of the current point cloud) and actively selects points that are worth intervening on. BRM then performs constrained residual refinement on the selected regions by aggregating hybrid spatial-feature context from both spatial and feature neighborhoods. The refined predictions induce state transitions, allowing AIA and BRM to jointly model prediction optimization as a multi-step Markov Decision Process (MDP).

For safeguarding reliable predictions, CRS introduces a high-confidence suppression reward, which penalizes unnecessary interventions on high-confidence predictions. Meanwhile, to prevent an overly conservative policy that trivially reduces interventions to avoid penalties, CRS incorporates a low-confidence exploration reward, which encourages interventions on low-confidence predictions to explore potentially erroneous regions. Experimental results on S3DIS and ScanNetV2 show that PARI achieves strong performance across multiple weak supervision settings with robust cross-backbone generalization, and achieves performance comparable to the reported fully supervised reference using the same backbone with only 1% annotations, highlighting its effectiveness under sparse supervision.

In summary, the main contributions of this work are as follows: *(i)* We propose a novel PARI framework for 3D-WSSS, which shifts the passive paradigm toward active intervention, thereby

explicitly resolving the lack of targeted intervention. *(ii)* We design a collaborative intervention mechanism that decomposes selective refinement into two coupled subproblems: AIA determines *where to intervene* by localizing high-risk points from uncertainty maps and prediction priors, while BRM determines *how to intervene* by performing residual refinement with hybrid spatial-feature context. *(iii)* We formulate prediction optimization as a multi-step MDP guided by a CRS strategy, which balances exploration and preservation by rewarding interventions on uncertain regions while suppressing unnecessary changes to reliable predictions.

2 Related Work

Weakly Supervised Point Cloud Semantic Segmentation reduces the need for exhaustive point-level annotations, benefiting downstream applications such as robotic manipulation [4, 11, 29] and perception [48, 19, 3, 50]. Recent advances can be broadly categorized into three paradigms: consistency regularization [43, 16, 56], pseudo-label propagation [6, 32, 7], and other emerging approaches [23, 52, 22, 33]. In pseudo-label propagation, Q2E [6] designs an image-assisted generator that uses 2D visual priors and a hierarchical optimization strategy to progressively expand and refine 3D annotations. Similarly, OTOC [20] employs a relation network to extract class prototypes and explicitly models graph node similarities to produce high-quality pseudo-labels. However, these methods suffer from *a lack of targeted intervention*, as erroneous predictions are neither explicitly identified nor corrected, leading to the continuous propagation of incorrect semantics and severe error accumulation, particularly near semantic boundaries and in geometrically complex regions. In consistency regularization, RAC-Net [43] assesses pseudo-label reliability using prediction confidence and model uncertainty, and imposes differentiated consistency constraints on reliable and ambiguous points. HybridCR [16] further captures feature correlations by combining standard segmentation objectives with auxiliary point-level consistency and contrastive constraints. Nevertheless, these approaches exhibit *limited contextual awareness*, as supervision is restricted to point- or view-level self-augmentation consistency, failing to capture the rich contextual dependencies inherent in complex scenes. Beyond these dominant paradigms, several works explore alternative perspectives for partially labeled scenarios. For example, AADNet [23] introduces a label-aware downsampling strategy to preserve annotated points and enhance supervision density during feature extraction. C3 [52] proposes a cross-scene contrastive regularization scheme that pulls prototypes of the same class closer while pushing different ones apart to maintain clear decision boundaries. However, despite these efforts, existing methods remain within a passive learning paradigm. Consequently, the fundamental challenges of error accumulation and limited contextual awareness persist, motivating our shift toward a Policy-driven Active Residual Intervention (PARI) framework.

Reinforcement Learning (RL) shows great promise for active intervention by leveraging reward signals to precisely identify high-risk points. Evolving from classic algorithms like Q-Learning [40] and REINFORCE [41], advanced RL has recently achieved remarkable progress in RL-based post-training of large language models: PPO [28] provides a practical balance between efficiency and stability via clipped surrogate objectives, DPO [27] efficiently aligns preferences without auxiliary reward models, and GRPO [30] scales complex reasoning by bypassing memory-intensive critics. Despite their success in LLMs, DPO and GRPO are not well-suited for 3D-WSSS. DPO directly optimizes the policy based on preference pairs, which is difficult to obtain under extremely sparse point-level annotations. In our 3D-WSSS setting, GRPO has been empirically shown to fail in producing discriminative group advantages, making it unable to distinguish beneficial actions. Consequently, we adopt PPO to optimize the Active Intervention Agent (AIA). Furthermore, PARI fundamentally differs from RL post-training: it formulates prediction refinement as a multi-step MDP rather than response-level generation or preference alignment, and employs an external AIA as the policy instead of directly optimizing the segmentation backbone itself.

3 Method

3.1 Overview

PARI follows a two-stage training pipeline, as shown in Fig. 2. In Stage I, PARI trains a base segmentation network with sparse point-level annotations to extract point-wise feature embeddings F and generate initial logits $Z^{(0)}$. Conditioned on feature embeddings and initial logits, PARI formulates

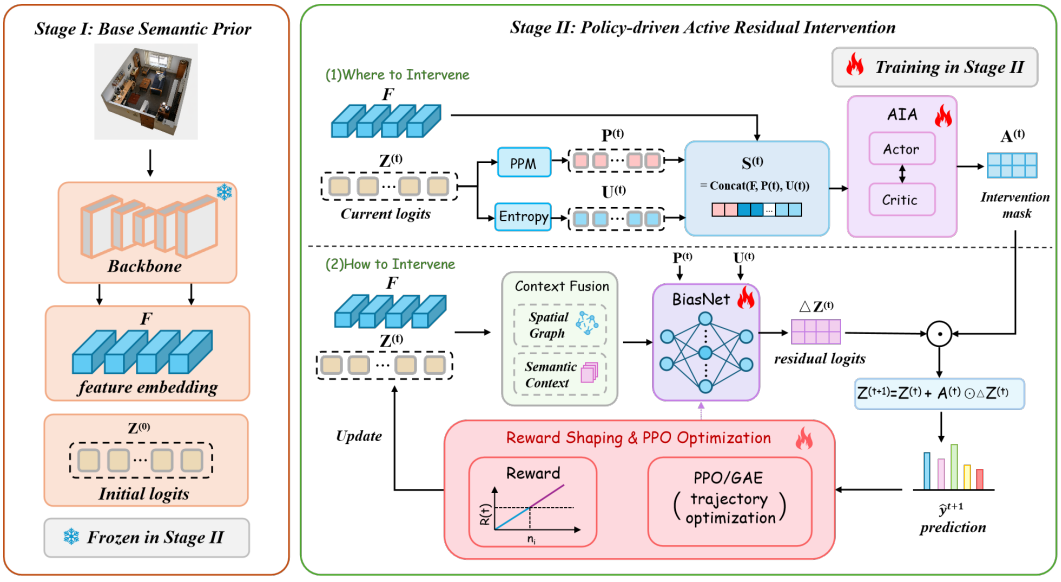

Figure 2: **Overview of the two-stage training pipeline for the proposed PARI framework.** Stage I learns a stable but imperfect semantic prior using a base segmentation network, producing point-wise semantic features and initial logits. Based on this frozen prior, Stage II performs multi-step active intervention over each batch of point clouds.

Stage II as a multi-step active intervention process, where at each step, an Active Intervention Agent (AIA) guided by a Confidence-aware Reward Shaping (CRS) strategy actively localizes high-risk points, and the Bias-aware Refinement Module (BRM) subsequently refines the predictions of these explicitly identified targets. At step t, AIA processes feature embeddings F, a multi-scale prediction prior $P^{(t)}$, and an uncertainty map $U^{(t)}$ to output a binary intervention mask $A^{(t)}$ that explicitly localizes high-risk predictions. This point-wise selection is guided by the CRS strategy, which utilizes an exploration reward to promote intervention on low-confidence areas and a suppression penalty to discourage unnecessary modifications to high-confidence regions. BiasNet, the residual prediction component of BRM, leverages the hybrid spatial-feature context aggregated from spatial and feature graphs, combined with the multi-scale prediction prior and an uncertainty map, to perform residual refinement on the high-risk points, yielding the updated logits $Z^{(t+1)}$ for the next step.

3.2 Active Intervention Agent: Where to Intervene

Given point coordinates $\mathcal{P} \in \mathbb{R}^{N\times 3}$ and input features $\mathcal{X} \in \mathbb{R}^{N\times d_x}$, the Stage-I segmentation network f_θ produces point-wise semantic features $F = f_\theta^{\text{emb}}(\mathcal{P}, \mathcal{X}) \in \mathbb{R}^{N\times D}$ and initial logits $Z^{(0)} = f_\theta^{\text{logit}}(\mathcal{P}, \mathcal{X}) \in \mathbb{R}^{N\times C}$, where D is the feature dimension and C is the number of semantic classes. In Stage II, PARI freezes f_θ and uses F and $Z^{(0)}$ as a frozen semantic prior for subsequent multi-step active intervention.

To explicitly localize high-risk points, PARI introduces an Active Intervention Agent (AIA), which formulates point-wise intervention selection as a multi-step Markov decision process. At each step, AIA observes a state composed of frozen semantic features, an uncertainty map, and multi-scale prediction priors, and dynamically predicts a point-wise intervention mask for subsequent residual refinement. Driven by multi-step reward feedback, AIA learns an adaptive intervention policy that prioritizes points whose updates are expected to improve future prediction states, while simultaneously identifying high-risk regions and preserving reliable predictions.

State Representation. At intervention step t, AIA observes the frozen semantic prior and the current prediction state. The frozen point-wise feature F provides stable semantic cues, while the current logits $Z^{(t)}$ encode predictions after previous interventions. From $Z^{(t)}$, we compute the class

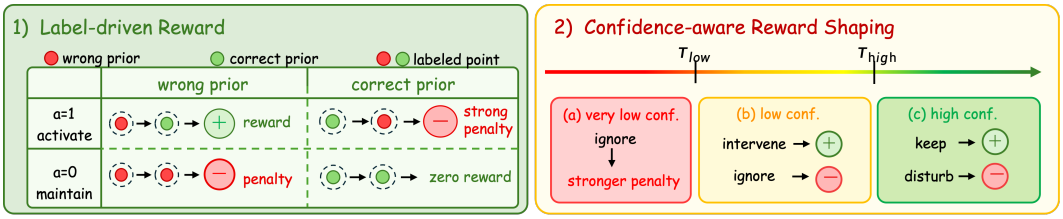

Figure 3: **Reward design with confidence-aware reward shaping.** Label-driven reward encourages fixing wrong priors and preserving correct priors, while confidence-aware shaping rewards low-confidence interventions and suppresses high-confidence activations. Appendix E.1 analyzes how these shaping signals affect PPO advantage estimation and policy updates.

probability map $\Pi^{(t)} = \mathrm{softmax}(Z^{(t)})$ and a normalized uncertainty map:

$$U_i^{(t)} = -\frac{1}{\log C} \sum_{c=1}^{C} \Pi_{i,c}^{(t)} \log \left(\Pi_{i,c}^{(t)} + \epsilon\right). \tag{1}$$

To complement point-wise uncertainty with local and multi-scale prediction cues, we adopt a lightweight Pyramid Pooling Module (PPM) to extract a multi-scale prediction prior $P^{(t)}$ from the current prediction map, with details provided in Appendix A.2. The final state representation is obtained by concatenating the frozen semantic feature, prediction prior, and uncertainty map:

$$S^{(t)} = \mathrm{Concat}(F, P^{(t)}, U^{(t)}). \tag{2}$$

Actor-Critic Policy. Given the state representation $S^{(t)}$, AIA adopts an actor-critic architecture for point-wise intervention decisions. The actor π_ω predicts an intervention probability for each point and samples a binary action:

$$a_i^{(t)} \sim \pi_\omega(\cdot \mid s_i^{(t)}), \quad a_i^{(t)} = \begin{cases} 1, & i \in \mathcal{I}^{(t)}, \\ 0, & i \notin \mathcal{I}^{(t)}, \end{cases} \tag{3}$$

where $\mathcal{I}^{(t)}$ denotes the set of points selected for intervention at step t. Aggregating all point-wise actions gives the intervention mask $A^{(t)}$, which determines where residual refinement is applied: selected points are sent to the Bias-aware Refinement Module (BRM), while the remaining points keep their current predictions. The critic V_ψ estimates the state value $V_\psi(S^{(t)})$ for advantage estimation in PPO. Through actor-critic learning, AIA dynamically allocates interventions based on reward feedback.

Reward Design with Confidence-aware Shaping. Under extremely sparse supervision, reward design is crucial for policy learning: sparse labeled feedback may lead to a lazy policy that avoids intervention, while unconstrained activation may corrupt reliable predictions. We therefore design a reward function that encourages interventions on potentially erroneous points and penalizes unnecessary activation on reliable ones. Formally, the reward consists of a label-driven feedback term and a confidence-aware shaping term:

$$r_i^{(t)} = r_{i,\mathrm{lab}}^{(t)} + r_{i,\mathrm{conf}}^{(t)}, \tag{4}$$

where $r_{i,\mathrm{lab}}^{(t)}$ provides supervised feedback on labeled points, and $r_{i,\mathrm{conf}}^{(t)}$ provides dense shaping signals on unlabeled points.

Label-driven reward. For labeled points, we define a label-driven reward according to the prediction correctness before and after intervention, together with the intervention action:

$$r_{i,\mathrm{lab}}^{(t)} = \mathcal{R}_{\mathrm{lab}}(\hat{y}_i^{(t)}, \hat{y}_i^{(t+1)}, y_i, a_i^{(t)}), \tag{5}$$

where $\hat{y}_i^{(t)}$ and $\hat{y}_i^{(t+1)}$ denote the predicted classes before and after intervention, y_i is the ground-truth label, and $a_i^{(t)}$ is the intervention action. As shown in Fig. 3 left, correcting a wrong prediction is rewarded, while missing a labeled error is penalized; conversely, preserving a correct prediction incurs no penalty, while making it incorrect incurs a penalty. This encourages the policy to correct unreliable priors while preserving reliable ones.

Confidence-aware reward shaping. Since labeled rewards remain sparse, we further exploit unlabeled points through confidence-aware reward shaping:

$$r_{i,\text{conf}}^{(t)} = \mathcal{R}_{\text{conf}}(c_i^{(t)}, a_i^{(t)}), \tag{6}$$

where $c_i^{(t)}$ is the prediction confidence and $a_i^{(t)}$ is the intervention action. As illustrated in Fig. 3 right, we use confidence thresholds τ_{low} and τ_{high} to define confidence-aware reward shaping for unlabeled points. High-confidence points are treated as reliable priors and are protected by penalizing unnecessary activation; low-confidence points are encouraged for intervention; and extremely low-confidence points receive stronger penalties when ignored. The empirical relationship between prediction confidence and reliability is analyzed in Appendix D.4 and D.5, where low-confidence intervals exhibit higher weak-region density and attract more interventions from AIA. This design converts unlabeled regions into weak policy feedback, encouraging useful interventions while suppressing destructive modifications.

State Transition. After sampling the intervention mask $A^{(t)}$, the selected points are sent to BiasNet for residual refinement, producing updated logits $Z^{(t+1)}$. The next state $S^{(t+1)}$ is then reconstructed from the updated logits, together with the corresponding uncertainty map and multi-scale prediction prior. This completes the MDP transition from the current state to the next, allowing the policy to progressively update predictions through multi-step active intervention.

3.3 Bias-aware Refinement Module: How to Intervene

Given the high-risk points localized by the intervention mask $A^{(t)}$ from AIA, PARI employs a Bias-aware Refinement Module (BRM) to conduct targeted refinement by exploiting rich spatial and semantic scene contexts. Within BRM, BiasNet serves as the residual prediction component, leveraging hybrid spatial-feature context aggregated from spatial and feature graphs, together with the multi-scale prediction prior and uncertainty map, to execute selective refinement.

Hybrid Spatial-Feature Context. First, BiasNet builds two complementary affinity graphs: a spatial graph $\mathcal{G}^s$ based on geometric proximity, and a feature graph $\mathcal{G}^f$ based on semantic similarity in the frozen prior feature space. Their neighborhood supports are obtained by affinity ranking:

$$\mathcal{N}^s(i) = \text{Top}_{K_s}^j\big(\rho_s(p_i, p_j)\big), \qquad \mathcal{N}^f(i) = \text{Top}_{K_f}^j\big(\rho_f(F_i, F_j)\big), \tag{7}$$

where ρ_s and ρ_f denote spatial and semantic affinities, respectively. The contexts aggregated from $\mathcal{G}^s$ and $\mathcal{G}^f$ are fused to combine local geometric consistency with non-local semantic support, providing BiasNet with richer contextual representations for residual prediction.

Attention-guided Residual Transition. Before residual prediction, BiasNet uses a lightweight attention mechanism to aggregate and fuse the spatial and feature contexts. The attention weights adaptively emphasize more informative neighbors and contextual cues, producing an enhanced representation for residual refinement. Following the framework in Fig. 2, BiasNet takes this enhanced representation, the multi-scale prediction prior $P^{(t)}$, and the uncertainty map $U^{(t)}$ as inputs, and predicts the residual logit offset $\Delta Z^{(t)}$ through a lightweight residual head. The prediction state is then updated through an action-gated residual transition:

$$Z^{(t+1)} = Z^{(t)} + A^{(t)} \odot \Delta Z^{(t)}, \tag{8}$$

where $\odot$ denotes point-wise gating broadcast along the class dimension. Here, $A^{(t)}$ specifies *where* residual refinement is applied, while $\Delta Z^{(t)}$ determines *how* the selected predictions are refined. The updated logits $Z^{(t+1)}$ are then used to construct the next decision state.

Joint Optimization and Output. AIA and BiasNet are jointly optimized through iterative intervention. At each episode, AIA samples intervention masks over T intervention steps, forming a trajectory $\tau = \{(S^{(t)}, A^{(t)}, r^{(t)}, S^{(t+1)})\}_{t=0}^{T-1}$. The intervention agent is optimized with the clipped PPO objective, which constrains policy updates and reinforces actions with higher estimated advantages:

$$\mathcal{L}_{\text{agent}} = -\mathbb{E}_t\Big[\min\Big(\rho_t(\omega)\hat{A}_t,\ \text{clip}(\rho_t(\omega), 1-\epsilon, 1+\epsilon)\hat{A}_t\Big)\Big] + \lambda_v \mathcal{L}_{\text{critic}} - \alpha\mathcal{H}(\pi_\omega), \tag{9}$$

where $\rho_t(\omega)$ is the policy ratio, $\hat{A}_t$ is the advantage estimate, $\mathcal{L}_{\text{critic}}$ is the value loss, and $\mathcal{H}(\pi_\omega)$ is the entropy regularizer. This objective encourages AIA to select actions that improve long-term refinement rewards while maintaining sufficient exploration.

Given the sampled intervention masks, BiasNet is optimized with the following multi-objective residual refinement loss, which jointly constrains semantic correction and prediction stability:

$$\mathcal{L}_{\text{bias}} = \mathcal{L}_{\text{CE}} + \lambda_1 \mathcal{L}_{\text{KL}} + \lambda_2 \mathcal{L}_{\text{ent}} + \lambda_3 \mathcal{L}_{\text{neigh}}. \tag{10}$$

Here, $\mathcal{L}_{\text{CE}}$ anchors residual refinement with sparse labels and high-confidence pseudo-labels, $\mathcal{L}_{\text{KL}}$ preserves reliable high-confidence regions, $\mathcal{L}_{\text{ent}}$ encourages uncertainty reduction on activated low-confidence points, and $\mathcal{L}_{\text{neigh}}$ improves local consistency with neighbor-induced targets. Detailed definitions and discussions of these loss terms are provided in Appendix A.4.

Table 1: Results on S3DIS under fully supervised and uniformly sparse annotation settings.

Setting	Method	mIoU	ceiling	floor	wall	beam	column	window	door	chair	table	bookcase	sofa	board	clutter
100% (Fully)	PointNet [24]	41.1	88.8	97.3	69.8	**0.1**	3.9	46.3	10.8	59.0	52.6	5.9	40.3	26.4	33.2
	SQN [9]	63.7	92.8	96.9	81.8	0.0	25.9	50.5	65.9	79.5	85.3	55.7	72.5	65.8	55.9
	HybridCR [16]	65.8	93.6	98.1	82.3	0.0	24.4	59.5	66.9	79.6	87.9	67.1	73.0	66.8	55.7
	ERDA [34]	68.3	93.9	98.5	83.4	0.0	28.9	**62.6**	70.0	**89.4**	82.7	75.5	69.5	75.3	58.7
	DeepGCN [15]	60.0	90.8	97.5	76.7	0.0	24.9	51.4	52.7	76.7	83.0	61.1	62.2	58.5	44.6
	PointNeXt [26]	69.2	**94.7**	98.5	82.9	0.0	24.2	59.9	74.3	83.0	**91.4**	76.3	75.5	**78.6**	**60.4**
	AADNet* [23]	69.3	94.3	**98.6**	82.6	0.0	31.3	57.1	**77.0**	81.0	90.6	**79.5**	73.9	74.8	59.5
1%	PointCT [36]	67.6	-	-	-	-	-	-	-	-	-	-	-	-	-
	DiCoSeg [17]	68.3	-	-	-	-	-	-	-	-	-	-	-	-	-
	AADNet [23]	68.3	-	-	-	-	-	-	-	-	-	-	-	-	-
	PARI (Ours)	**69.4**	92.9	98.4	**84.4**	0.0	**36.3**	57.8	72.3	82.7	90.0	79.3	**76.4**	75.9	55.9
0.1%	SQN [9]	61.4	91.7	95.6	78.7	0.0	24.2	55.9	63.1	70.5	83.1	60.7	67.8	56.1	50.6
	CPCM [18]	66.3	91.4	95.5	82.0	0.0	**30.8**	54.1	70.1	**87.6**	79.4	70.0	67.0	**77.8**	**56.6**
	PointMatch [42]	63.4	-	-	-	-	-	-	-	-	-	-	-	-	-
	DeepGCN [15]	43.9	-	-	-	-	-	-	-	-	-	-	-	-	-
	PointNeXt [26]	65.0	**93.7**	97.8	79.5	0.0	27.3	59.2	62.2	79.4	**88.6**	64.2	70.1	69.3	53.6
	AADNet [23]	67.2	**93.7**	**98.0**	81.5	0.0	19.4	59.5	72.0	80.9	88.5	78.3	73.0	72.1	56.1
	PARI (Ours)	**67.9**	93.3	97.8	**82.3**	0.0	29.4	**59.9**	**72.3**	79.5	88.2	**79.9**	**74.1**	69.9	55.7
0.03%	PSD [54]	48.2	87.9	96.0	62.1	0.0	20.6	49.3	40.9	55.1	61.9	43.9	50.7	27.3	31.1
0.03%	HybridCR [16]	51.5	85.4	91.9	65.9	0.0	18.0	51.4	34.2	63.8	78.3	52.4	59.6	29.9	39.0
0.03%	DCL [49]	59.6	91.7	95.8	76.4	0.0	21.2	**58.3**	29.6	72.6	83.3	64.2	**69.6**	**63.4**	48.6
0.02%	MILTrans [47]	51.4	86.6	93.2	75.0	0.0	**29.3**	45.3	46.7	60.5	62.3	56.5	47.5	33.7	32.2
0.02%	ERDA [34]	48.4	87.3	96.3	61.9	0.0	11.3	45.9	31.7	73.1	65.1	57.8	26.1	36.0	36.4
0.02%	MulPro [32]	47.5	90.1	96.3	71.8	0.0	6.7	46.7	39.2	67.2	67.4	21.8	39.2	33.0	38.0
0.01%	SQN [9]	45.3	89.2	93.5	71.3	0.0	4.1	34.7	41.0	54.9	66.9	25.7	55.4	12.8	39.6
0.01%	CPCM [18]	59.3	-	-	-	-	-	-	-	-	-	-	-	-	-
0.01%	DeepGCN [15]	35.9	-	-	-	-	-	-	-	-	-	-	-	-	-
0.01%	PointNeXt [26]	58.4	89.4	96.5	75.7	**0.1**	22.6	55.3	44.5	**74.2**	84.3	54.2	62.8	52.5	47.0
0.01%	AADNet [23]	60.8	**92.5**	96.6	77.2	0.0	20.9	57.0	**61.1**	72.2	83.1	60.1	67.8	52.9	**49.0**
0.01%	**PARI (Ours)**	**61.4**	90.8	**97.0**	**78.8**	**0.1**	24.9	54.3	58.5	72.6	**84.6**	**66.9**	68.2	56.0	45.7

4 Experiments

4.1 Experimental Setting

Datasets and metrics. We evaluate PARI on S3DIS [1] and ScanNetV2 [5]. S3DIS contains 271 rooms from six indoor areas with 13 categories, and we use Area 5 for validation following the common protocol. ScanNetV2 contains 1,513 scans from 707 indoor environments with 20 categories, and we follow the official 1,201/312 train/validation split. We use mean Intersection-over-Union (mIoU) as the primary metric. SemanticKITTI [2] experiments are reported in Appendix D.6 for outdoor transfer evaluation.

Implementation details. We use AADNet [23] as the default backbone unless otherwise specified, and compare all methods under the same preprocessing, sampling strategy, and sparse annotation protocol. We evaluate uniform annotations on S3DIS with 1%, 0.1%, and 0.01% annotations, and on ScanNetV2 with 1% annotations and the 20pts setting; non-uniform S3DIS annotations and alternative backbones are further tested for robustness and generality. PARI runs T intervention steps on a frozen base semantic prior, with AIA optimized by PPO and BRM optimized by the refinement loss in Sec. 3.3. More implementation details are provided in Appendix B.

4.2 Main Results

Results on S3DIS under uniform sparse annotations. To evaluate PARI under varying sparsity levels, we compare it with fully supervised and weakly supervised methods on S3DIS using uniform sparse annotations, as reported in Table 1. Most baseline results are taken from DGNet [22]. Here, **bold** and underline indicate the best and second-best performance, respectively. Under the 1% annotation setting, PARI achieves 69.4 mIoU, compared with 68.3 for AADNet, and performs comparably to the reported fully supervised reference using the same backbone. With 0.1% and 0.01% annotations, PARI achieves 67.9 and 61.4 mIoU, respectively, exceeding the AADNet results reported in Table 1 by 0.7 and 0.6 percentage points. These results support the effectiveness of selective residual intervention across the evaluated annotation ratios. By selectively intervening on high-risk predictions while preserving reliable ones, PARI combines sparse labeled feedback with confidence-aware signals from unlabeled points to guide targeted refinement of the base prior.

Results on ScanNetV2 under sparse annotations. To analyze PARI on more challenging scenes and settings, we compare it with existing methods on ScanNetV2, as shown in Table 2, where * denotes our reproduced results. Under 1% annotations, PARI achieves 67.2 mIoU, comparable to the reported fully supervised reference using the same backbone. Under the extreme 20pts setting, PARI still obtains 60.8 mIoU and outperforms most existing methods. These results show that PARI remains effective on complex indoor scenes, even with extremely sparse supervision.

Results under non-uniform annotations on S3DIS. To evaluate robustness to biased supervision, we test PARI on S3DIS with non-uniform annotations, as reported in Table 3. We use 10% and 1% annotations, since lower ratios such as 0.1% or 0.01% would leave many regions or categories nearly unlabeled, making the evaluation dominated by missing labels rather than annotation imbalance. PARI achieves the best mIoU under both settings, reaching 57.40 and 38.55 mIoU, with gains of 0.34 and 0.61 points over the paired AADNet baselines, respectively. They show that PARI remains robust under imbalanced supervision and can refine unreliable predictions even with a biased semantic prior.

Generalization across different backbones. To assess architectural generality, we integrate PARI with RandLA-Net [10] and AADNet on S3DIS under the same sparse supervision protocol. Table 4 shows consistent gains on both backbones: +2.2/+0.6 mIoU for RandLA-Net and +1.1/+0.7 mIoU for AADNet under 1%/0.1% annotations. This confirms that PARI acts as a general intervention-and-refinement framework rather than a backbone-specific design.

Table 2: Results on ScanNetV2 under sparse annotation settings.

Setting	Method	mIoU (%)
100% (Fully)	PointNet [24]	33.9
	PointNet++ [25]	53.5
	HybridCR [16]	59.9
	AADNet [23]	66.9
1%	Zhang *et al.* [53]	51.1
	PSD [54]	54.7
	HybridCR [16]	56.8
	DCL [49]	59.6
	GaIA [13]	65.2
	ERDA [34]	63.0
	AADNet* [23]	<u>66.8</u>
	PARI (Ours)	**67.2**
20pts	Hou *et al.* [8]	55.5
	OTOC [20]	59.4
	MILTrans [47]	54.4
	DAT [44]	55.2
	PointMatch [42]	**62.4**
	ERDA [34]	57.0
	AADNet* [23]	60.2
	PARI (Ours)	<u>60.8</u>

Table 3: Results under non-uniform annotation distribution on S3DIS.

Setting	Method	mIoU (%)
10%	DCL [49]	52.9
	CPCM [18]	54.9
	PointNeXt [26]	54.1
	AADNet* [23]	57.1
	PARI (Ours)	**57.4**
1%	DCL [49]	35.6
	CPCM [18]	38.2
	PointNeXt [26]	35.9
	AADNet* [23]	37.9
	PARI (Ours)	**38.6**

Table 4: Backbone generalization results on S3DIS under 1% and 0.1% uniform sparse annotations. Gains brought by PARI are highlighted for different backbone networks.

Backbone	1% annotations			0.1% annotations		
	Baseline	**+PARI**	**Gain**	**Baseline**	**+PARI**	**Gain**
RandLA-Net [10]	42.4	**44.6**	**+2.2**	43.2	**43.8**	**+0.6**
AADNet [23]	68.3	**69.4**	**+1.1**	67.2	**67.9**	**+0.7**

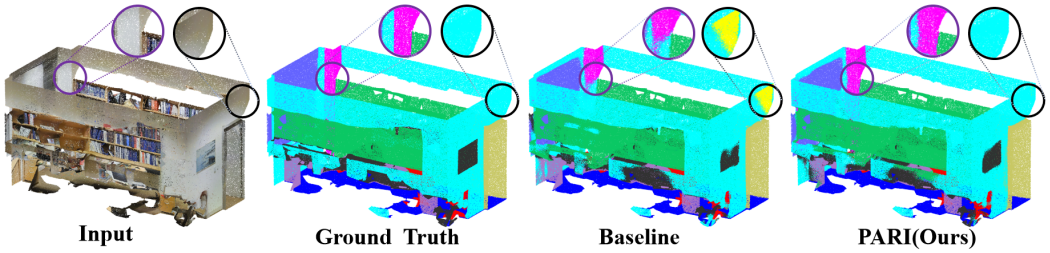

Figure 4: Qualitative visualization on S3DIS. From left to right: input point cloud, ground truth, baseline prediction, and PARI prediction.

Table 5: Ablation studies on S3DIS under uniform sparse annotations. (a) Module ablation. (b) Reward design ablation. (c) Loss design ablation.

Variant	BRM	AIA	1%	0.1%	0.01%
Base model	×	×	68.3	67.2	60.8
+ BRM	✓	×	68.9	66.9	60.4
+ BRM + AIA	✓	✓	**69.4**	**67.9**	**61.4**

(a) Module ablation

R_{sup}	R_{low}	R_{high}	1%	0.1%	0.01%
✓			69.1	66.7	60.8
✓	✓		69.3	67.3	60.6
✓		✓	68.7	67.3	61.0
✓	✓	✓	**69.4**	**67.9**	**61.4**

(b) Reward design ablation

Setting	1%	0.1%	0.01%
w/o Ent	69.3	67.4	61.3
w/o KL	68.4	67.3	61.0
w/o Neigh	69.0	66.5	60.9
Full	**69.4**	**67.9**	**61.4**

(c) Loss design ablation

Qualitative visualization. To further examine the effect of PARI, we visualize segmentation results on S3DIS in Fig. 4. Compared with the baseline, PARI produces cleaner predictions that are more consistent with the ground truth, especially around object boundaries, small objects, and semantically complex regions. In the left enlarged region, PARI produces a clearer semantic boundary that better matches the ground truth. In the right enlarged region, it reduces misclassified points on the wall, yielding more consistent predictions. This suggests that policy-driven active residual intervention can better localize unreliable predictions and refine them selectively.

4.3 Ablation Studies

Ablation study for core modules. To verify the necessity of policy selection and residual refinement, Table 5(a) ablates the core modules of PARI. Using residual refinement alone improves performance under the 1% setting, but degrades mIoU under 0.1% and 0.01% annotations, suggesting that residual updates without explicit selection may disturb reliable semantic priors when supervision is extremely sparse. After introducing AIA, PARI improves consistently across all annotation ratios. These results indicate that residual refinement alone is insufficient for stable gains, and explicitly modeling *where to intervene* is critical before applying residual updates.

Ablation study for reward design.

To analyze the role of each reward term, Table 5(b) compares supervised reward only, w/o low-confidence exploration reward, w/o high-confidence activation penalty, and the full reward design. The results show that supervised reward alone provides limited guidance for unlabeled regions, while removing either confidence-aware term leads to performance drops. This indicates that the low-confidence exploration reward and high-confidence activation penalty are complementary: the former promotes exploration of uncertain regions, while the latter suppresses unnecessary intervention on reliable predictions. Together, they balance effective intervention and reliable preservation, enabling the full reward design to achieve the best performance.

Ablation study for refinement losses. To evaluate the role of each BiasNet loss term, Table 5(c) removes the entropy, KL, and neighbor consistency losses, respectively. The results show that removing any term degrades performance. This indicates that these losses stabilize residual refinement

Table 6: Intervention analysis on S3DIS (1% labels): (a) local budgets; (b) selection at a 10% local budget. Fix–Break: corrected minus newly introduced errors.

Policy	Budget	mIoU
Actor top-k	5%	69.307
Actor top-k	10%	69.357
Actor top-k	20%	69.395
Adaptive (default)	None	69.402

(a) Intervention budget

Selection	mIoU	Fix–Break
Random	69.254	+3,384
Entropy top-k	69.323	+4,757
Actor top-k	**69.357**	**+21,792**

(b) Selection at a 10% budget

Table 7: Training and inference efficiency. Memory denotes peak allocated GPU memory.

Training metric	AADNet Stage I	PARI Stage II
Time per epoch (min)	13.39	13.42
Memory (MiB)	8,667	4,638
Inference metric	**AADNet**	**AADNet + PARI**
Latency (ms)	64.32	67.24
FLOPs (G)	24.40	24.76
Memory (MiB)	240.80	250.66
Parameters (M)	7.1251	7.1340

from complementary aspects: reducing uncertainty, constraining semantic drift, and preserving local consistency, enabling more stable updates under sparse supervision.

Intervention budget and point-selection strategies. We examine budget sensitivity and learned point selection on S3DIS with 1% annotations. At inference, the budget specifies the fraction of highest-scoring points selected for intervention within each test sub-cloud at each intervention step. As shown in Table 6(a), increasing this fraction from 5% to 20% changes mIoU only from 69.307 to 69.395, indicating low sensitivity within this range. Both results are close to the default adaptive policy's 69.402 mIoU, which imposes no fixed selection ratio. To compare selection strategies, we further evaluate Random, Entropy top-k, and Actor top-k under the same 10% local budget. Table 6(b) shows that Actor selection achieves the highest mIoU and the largest number of net corrections. Although its mIoU advantage over entropy selection is modest, the increase from 4,757 to 21,792 net corrections further indicates that the learned policy selects more beneficial intervention locations.

4.4 Computational Efficiency

We evaluate the computational cost of PARI on a single NVIDIA A100 GPU. Training time is averaged over ten complete epochs, while inference is measured using the S3DIS 1% checkpoint on 24k-point test sub-clouds. As shown in Table 7, Stage II takes 13.42 minutes per epoch, compared with 13.39 minutes for Stage I, and reduces peak allocated training memory from 8,667 to 4,638 MiB. Each Stage-II batch requires only one forward pass through the frozen backbone, with cached features reused across two intervention steps. The Stage-I checkpoint is trained for 49 epochs, and Stage II is configured for 100 epochs; the complete training cost includes both stages. At inference, a single intervention step increases model-forward latency from 64.32 to 67.24 ms (+4.54%) and adds only 0.0089 M parameters (+0.125%), enabling selective refinement with modest inference overhead.

5 Conclusion

We presented PARI, a policy-driven active residual intervention framework for weakly supervised point cloud semantic segmentation. PARI explicitly models *where to intervene* and *how to intervene*: AIA selects high-risk points, while BRM applies constrained residual refinement. Confidence-aware reward shaping stabilizes policy learning by exploiting unlabeled points as dense feedback under extreme annotation sparsity. Experiments demonstrate the effectiveness, robustness, and backbone generality of PARI across the evaluated annotation settings. The ablation studies further show that residual refinement alone is insufficient for consistent gains, highlighting the importance of learned intervention selection. Decoupling intervention learning from backbone training makes PARI potentially complementary to existing weakly supervised methods.

Acknowledgments and Disclosure of Funding

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 62476260, 62225207 and 62436008, the Fundamental Research Funds for the Central Universities under Grant WK2100000057, and the CCF-1688 Yuanbao Cooperation Fund under Grant No. 2025006.

References

[1] Iro Armeni, Ozan Sener, Amir R Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 3d semantic parsing of large-scale indoor spaces. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1534–1543, 2016.

[2] Jens Behley, Martin Garbade, Andres Milioto, Jan Quenzel, Sven Behnke, Cyrill Stachniss, and Jurgen Gall. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9297–9307, 2019.

[3] Jingdao Chen, Yong Kwon Cho, and Zsolt Kira. Multi-view incremental segmentation of 3-d point clouds for mobile robots. *IEEE Robotics and Automation Letters*, 4(2):1240–1246, 2019.

[4] Eugenio Chisari, Nick Heppert, Max Argus, Tim Welschehold, Thomas Brox, and Abhinav Valada. Learning robotic manipulation policies from point clouds with conditional flow matching. In *Conference on Robot Learning*, pages 982–993. PMLR, 2025.

[5] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5828–5839, 2017.

[6] Jiacheng Deng, Jiahao Lu, and Tianzhu Zhang. Quantity-quality enhanced self-training network for weakly supervised point cloud semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(5):3580–3596, 2025.

[7] Anan Du, Tianfei Zhou, Shuchao Pang, Qiang Wu, and Jian Zhang. Pcl: Point contrast and labeling for weakly supervised point cloud semantic segmentation. *IEEE Transactions on Multimedia*, 26:8902–8914, 2024.

[8] Ji Hou, Benjamin Graham, Matthias Nießner, and Saining Xie. Exploring data-efficient 3d scene understanding with contrastive scene contexts. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 15587–15597, 2021.

[9] Qingyong Hu, Bo Yang, Guangchi Fang, Yulan Guo, Aleš Leonardis, Niki Trigoni, and Andrew Markham. Sqn: Weakly-supervised semantic segmentation of large-scale 3d point clouds. In *European conference on computer vision*, pages 600–619. Springer, 2022.

[10] Qingyong Hu, Bo Yang, Linhai Xie, Stefano Rosa, Yulan Guo, Zhihua Wang, Niki Trigoni, and Andrew Markham. Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11108–11117, 2020.

[11] David Inkyu Kim and Gaurav S Sukhatme. Semantic labeling of 3d point clouds with object affordance for robot manipulation. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5578–5584. IEEE, 2014.

[12] Roman Klokov and Victor Lempitsky. Escape from cells: Deep kd-networks for the recognition of 3d point cloud models. In *Proceedings of the IEEE international conference on computer vision*, pages 863–872, 2017.

[13] Min Seok Lee, Seok Woo Yang, and Sung Won Han. Gaia: Graphical information gain based attention network for weakly supervised point cloud semantic segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 582–591, 2023.

[14] Buyu Li, Wanli Ouyang, Lu Sheng, Xingyu Zeng, and Xiaogang Wang. Gs3d: An efficient 3d object detection framework for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1019–1028, 2019.

[15] Guohao Li, Matthias Muller, Ali Thabet, and Bernard Ghanem. Deepgcns: Can gcns go as deep as cnns? In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9267–9276, 2019.

[16] Mengtian Li, Yuan Xie, Yunhang Shen, Bo Ke, Ruizhi Qiao, Bo Ren, Shaohui Lin, and Lizhuang Ma. Hybridcr: Weakly-supervised 3d point cloud semantic segmentation via hybrid contrastive regularization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14930–14939, 2022.

[17] Jiawei Lian, Zhengxue Wang, Wentao Qu, Haobo Jiang, Le Hui, and Jian Yang. Diffusion-based contextual reconstruction for point cloud segmentation with limited annotations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 6780–6788, 2026.

[18] Lizhao Liu, Zhuangwei Zhuang, Shangxin Huang, Xunlong Xiao, Tianhang Xiang, Cen Chen, Jingdong Wang, and Mingkui Tan. Cpcm: Contextual point cloud modeling for weakly-supervised point cloud semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 18413–18422, 2023.

[19] Ming Liu. Robotic online path planning on point cloud. *IEEE transactions on cybernetics*, 46(5):1217–1228, 2015.

[20] Zhengzhe Liu, Xiaojuan Qi, and Chi-Wing Fu. One thing one click: A self-training approach for weakly supervised 3d semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1726–1736, 2021.

[21] Huimin Ma, Sheng Yi, Shijie Chen, Jiansheng Chen, and Yu Wang. Few annotated pixels and point cloud based weakly supervised semantic segmentation of driving scenes. *International Journal of Computer Vision*, 133(4):2096–2110, 2025.

[22] Zhiyi Pan, Wei Gao, Shan Liu, and Ge Li. Distribution guidance network for weakly supervised point cloud semantic segmentation. *Advances in neural information processing systems*, 37:32400–32420, 2024.

[23] Zhiyi Pan, Nan Zhang, Wei Gao, Shan Liu, and Ge Li. Point cloud semantic segmentation with sparse and inhomogeneous annotations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 6354–6362, 2025.

[24] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017.

[25] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, 30, 2017.

[26] Guocheng Qian, Yuchen Li, Houwen Peng, Jinjie Mai, Hasan Hammoud, Mohamed Elhoseiny, and Bernard Ghanem. Pointnext: Revisiting pointnet++ with improved training and scaling strategies. *Advances in neural information processing systems*, 35:23192–23204, 2022.

[27] Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36:53728–53741, 2023.

[28] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[29] Daniel Seita, Yufei Wang, Sarthak J Shetty, Edward Yao Li, Zackory Erickson, and David Held. Toolflownet: Robotic manipulation with tools via predicting tool flow from point clouds. In *Conference on Robot Learning*, pages 1038–1049. PMLR, 2023.

[30] Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

[31] Xibin Song, Peng Wang, Dingfu Zhou, Rui Zhu, Chenye Guan, Yuchao Dai, Hao Su, Hongdong Li, and Ruigang Yang. Apollocar3d: A large 3d car instance understanding benchmark for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5452–5462, 2019.

[32] Yongyi Su, Xun Xu, and Kui Jia. Weakly supervised 3d point cloud segmentation via multi-prototype learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(12):7723–7736, 2023.

[33] Tianfang Sun, Zhizhong Zhang, Xin Tan, Yanyun Qu, and Yuan Xie. Image understands point cloud: Weakly supervised 3d semantic segmentation via association learning. *IEEE Transactions on Image Processing*, 33:1838–1852, 2024.

[34] Liyao Tang, Zhe Chen, Shanshan Zhao, Chaoyue Wang, and Dacheng Tao. All points matter: Entropy-regularized distribution alignment for weakly-supervised 3d segmentation. *Advances in Neural Information Processing Systems*, 36:78657–78673, 2023.

[35] Sen Tao, Kaiduo Feng, Jiawei Liu, Peng Zeng, Yongchao Xu, Yufei Zheng, and Zheng-Jun Zha. PSP: Prompt-guided self-training sampling policy for active prompt learning. In *International Conference on Learning Representations*, 2026.

[36] Anh-Thuan Tran, Hoanh-Su Le, Suk-Hwan Lee, and Ki-Ryong Kwon. Pointct: Point central transformer network for weakly-supervised point cloud semantic segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3556–3565, 2024.

[37] Jingyi Wang, Yu Liu, Hanlin Tan, and Maojun Zhang. A survey on weakly supervised 3d point cloud semantic segmentation. *IET Computer Vision*, 18(3):329–342, 2024.

[38] Junfeng Wang, Jiawei Liu, Yongchao Xu, Tao Jiang, Jiangbo Ai, Jin Zhang, Chong Wang, and Qixing Zhang. Learning where to look: Observation policy optimization for thinking with images. *arXiv preprint arXiv:2611.04943*, 2026.

[39] Lei Wang, Yuchun Huang, Yaolin Hou, Shenman Zhang, and Jie Shan. Graph attention convolution for point cloud semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10296–10305, 2019.

[40] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3):279–292, 1992.

[41] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3):229–256, 1992.

[42] Yushuang Wu, Zizheng Yan, Shengcai Cai, Guanbin Li, Xiaoguang Han, and Shuguang Cui. Pointmatch: A consistency training framework for weakly supervised semantic segmentation of 3d point clouds. *Computers & Graphics*, 116:427–436, 2023.

[43] Zhonghua Wu, Yicheng Wu, Guosheng Lin, and Jianfei Cai. Reliability-adaptive consistency regularization for weakly-supervised point cloud segmentation. *International Journal of Computer Vision*, 132(6):2276–2289, 2024.

[44] Zhonghua Wu, Yicheng Wu, Guosheng Lin, Jianfei Cai, and Chen Qian. Dual adaptive transformations for weakly supervised point cloud segmentation. In *European conference on computer vision*, pages 78–96. Springer, 2022.

[45] Yongchao Xu, Jiawei Liu, Junfeng Wang, Sen Tao, Na Jiang, and Zheng-Jun Zha. Learning to diversify and focus: A reinforcement framework for open-vocabulary HOI detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 34673–34682, 2026.

[46] Yongchao Xu, Bowen Ye, Jiefeng Gan, Junkai Ma, Wenzhao Li, Sen Tao, Yi Wei, and Jiawei Liu. VideoEvolve: Co-evolving memory and retrieval for long video understanding. *arXiv preprint arXiv:2610.10183*, 2026.

[47] Cheng-Kun Yang, Ji-Jia Wu, Kai-Syun Chen, Yung-Yu Chuang, and Yen-Yu Lin. An mil-derived transformer for weakly supervised point cloud segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11830–11839, 2022.

[48] Lei Yang, Yanhong Liu, Jinzhu Peng, and Zize Liang. A novel system for off-line 3d seam extraction and path planning based on point cloud segmentation for arc welding robot. *Robotics and Computer-Integrated Manufacturing*, 64:101929, 2020.

[49] Baochen Yao, Hui Xiao, Jiayan Zhuang, and Chengbin Peng. Weakly supervised learning for point cloud semantic segmentation with dual teacher. *IEEE Robotics and Automation Letters*, 8(10):6347–6354, 2023.

[50] Anestis Zaganidis, Li Sun, Tom Duckett, and Grzegorz Cielniak. Integrating deep semantic segmentation into 3-d point cloud registration. *IEEE Robotics and automation letters*, 3(4):2942–2949, 2018.

[51] Fanrui Zhang, Qiang Zhang, Sizhuo Zhou, Jianwen Sun, Chuanhao Li, Jiaxin Ai, Yukang Feng, Yujie Zhang, Wenjie Li, Zizhen Li, Yifan Chang, Jiawei Liu, and Kaipeng Zhang. Code-in-the-loop forensics: Agentic tool use for image forgery detection. In *European Conference on Computer Vision*, volume 17023, pages 203–221, 2026.

[52] Yachao Zhang, Yuxiang Lan, Yuan Xie, Cuihua Li, and Yanyun Qu. Cross-cloud consistency for weakly supervised point cloud semantic segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, 36(8):14452–14463, 2025.

[53] Yachao Zhang, Zonghao Li, Yuan Xie, Yanyun Qu, Cuihua Li, and Tao Mei. Weakly supervised semantic segmentation for large-scale point cloud. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 3421–3429, 2021.

[54] Yachao Zhang, Yanyun Qu, Yuan Xie, Zonghao Li, Shanshan Zheng, and Cuihua Li. Perturbed self-distillation: Weakly supervised large-scale point cloud semantic segmentation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 15520–15528, 2021.

[55] Cong Zhao, Delong Ding, Cailin Lei, Shiyu Wang, Yuxiong Ji, and Yuchuan Du. Safety field-based vehicle-infrastructure cooperative perception for autonomous driving using 3d point clouds. *IEEE Transactions on Intelligent Transportation Systems*, 2025.

[56] Ce Zhou and Qiang Ling. Weakly supervised large-scale point cloud semantic segmentation based on dual consistency constraints and uncertainty-aware fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 63:1–13, 2024.

PARI: Policy-Driven Active Residual Intervention for Weakly Supervised Point Cloud Segmentation Appendix

Contents:

A Complete Method Details **16**
A.1 Overview of the PARI Pipeline 16
A.2 Multi-scale Prediction Prior Construction with PPM 16
A.3 AIA Architecture and Optimization 17
A.4 BRM and BiasNet Architecture 17
A.5 Detailed Reward Design 18
B Implementation Details **19**
C Code and Datasets **19**
C.1 Code 19
C.2 Datasets and Annotation Protocol Details 19
D Additional Experiments and Analysis **20**
D.1 Ablation Study for Top-k Neighborhood Size 20
D.2 Ablation Study for the Number of Intervention Steps 20
D.3 Ablation Study for Reward Coefficients 21
D.4 Confidence–Reliability Correlation on Unlabeled Predictions 21
D.5 Reward-guided Actor Focus Analysis 22
D.6 Additional Results on SemanticKITTI 23
D.7 Effect of High-confidence Pseudo-labels 23
D.8 Results across random seeds 23
D.9 Comparison with reproduced DGNet 23
D.10 Additional Qualitative Results 24
E Theoretical Analysis and Discussion **25**
E.1 How Confidence-aware Reward Shaping from Unlabeled Regions Affects Agent Decisions 25
E.2 How Policy-based Point Selection Improves Intervention Utility 25
E.3 How Selective Intervention Reduces Expected Prediction Error 26
E.4 Discussion: From Feedback-driven Post-training to Semantic Intervention 26
F Limitations and Future Work **27**

A Complete Method Details

A.1 Overview of the PARI Pipeline

Algorithm 1 summarizes the overall PARI pipeline. Stage I first learns a base segmentation model from sparse annotations and freezes it as the semantic prior for subsequent refinement. Stage II then performs multi-step active intervention, where AIA identifies points requiring further refinement and BiasNet updates only the selected predictions through residual refinement. The updated predictions are fed back to the next intervention step, forming a closed-loop refinement process that progressively improves the final dense prediction.

Algorithm 1 Policy-Driven Active Residual Intervention (PARI)

Input: Point cloud coordinates $\mathcal{P}$ and input features $\mathcal{X}$, intervention steps T
Given: Base segmentation network f_θ, AIA policy π_ω, BiasNet, PPM
Output: Final dense segmentation predictions

 # Stage I: Base Semantic Prior Learning
Optimize f_θ with sparse labels using cross-entropy loss
Extract feature embeddings $F = f_\theta^{emb}(\mathcal{P}, \mathcal{X})$ and initial logits $Z^{(0)} = f_\theta^{logit}(\mathcal{P}, \mathcal{X})$, then freeze f_θ

 # Stage II: Decoupled Active Residual Intervention
while training **do**
  Initialize trajectory buffer $\tau \leftarrow \emptyset$
  for $t = 0, 1, \ldots, T-1$ **do**
    # 1. State Modeling
    Compute probabilities $\Pi^{(t)} = \text{softmax}(Z^{(t)})$ and uncertainty map $U^{(t)}$
    Extract multi-scale prediction prior $P^{(t)} = \text{PPM}(Z^{(t)})$
    Formulate state $S^{(t)} = \text{Concat}(F, P^{(t)}, U^{(t)})$
    # 2. Where to Intervene (AIA)
    Sample binary intervention mask $A^{(t)} \sim \pi_\omega(\cdot \mid S^{(t)})$
    # 3. How to Intervene (BRM)
    Predict bounded residual logit offset $\Delta Z^{(t)}$ via BiasNet
    Update logits: $Z^{(t+1)} = Z^{(t)} + A^{(t)} \odot \Delta Z^{(t)}$
    # 4. Reward Calculation
    Compute step reward: $r^{(t)} = r_{lab}^{(t)} + r_{conf}^{(t)}$
    Store transition $(S^{(t)}, A^{(t)}, r^{(t)}, S^{(t+1)})$ in τ
  end for
  # Joint Optimization
  Compute Returns and Generalized Advantage Estimation (GAE) from τ
  Update AIA π_ω by minimizing the PPO loss $\mathcal{L}_{\text{agent}}$
  Update BiasNet by minimizing multi-objective loss $\mathcal{L}_{bias} = \mathcal{L}_{CE} + \lambda_1 \mathcal{L}_{KL} + \lambda_2 \mathcal{L}_{ent} + \lambda_3 \mathcal{L}_{neigh}$
end while
Return Final predictions: $\text{argmax}(Z^{(T)})$

A.2 Multi-scale Prediction Prior Construction with PPM

At each intervention step t, we use a Pyramid Pooling Module (PPM) to extract the multi-scale prediction prior $P^{(t)}$ from the current prediction map. Given logits $Z^{(t)}$, we first obtain the point-wise prediction map $\tilde{P}^{(t)}$ and feed it into PPM with pooling scales $\{1, 2, 3, 6\}$. The multi-scale features are then projected, upsampled, and fused to produce the multi-scale prediction prior:

$$P^{(t)} = \psi\left(\tilde{P}^{(t)}, \left\{\text{Up}\left(\phi_k\left(\text{Pool}_k(\tilde{P}^{(t)})\right)\right)\right\}_{k\in\{1,2,3,6\}}\right), \tag{11}$$

where Pool_k, ϕ_k, Up, and ψ denote scale-k pooling, branch projection, upsampling, and feature fusion, respectively. The resulting multi-scale prediction prior $P^{(t)}$ is concatenated with the frozen semantic feature F and uncertainty map $U^{(t)}$ as part of the AIA state.

A.3 AIA Architecture and Optimization

AIA learns a point-wise intervention policy to determine *where to intervene*. At intervention step t, it takes the state representation $S^{(t)} \in \mathbb{R}^{B\times N\times C_s}$ as input, where $C_s = D + C + 1$ corresponds to the frozen semantic feature, current prediction context, and uncertainty map. Both actor and critic are implemented as lightweight two-layer MLPs with a hidden dimension of 32. The actor predicts a binary action distribution for each point, while the critic estimates the corresponding state value for advantage estimation.

The policy and value function are defined as:

$$\pi_\omega(a_i^{(t)} \mid s_i^{(t)}) = \mathrm{Softmax}\big(f_\omega(s_i^{(t)})\big), \quad a_i^{(t)} \in \{0, 1\}, \tag{12}$$

$$V_\psi(s_i^{(t)}) = g_\psi(s_i^{(t)}), \tag{13}$$

where f_ω and g_ψ denote the actor and critic MLPs, respectively. Advantages are estimated with generalized advantage estimation (GAE) and normalized over valid points before PPO optimization.

AIA is optimized with the standard PPO objective:

$$\rho_t(\omega) = \frac{\pi_\omega(a^{(t)} \mid S^{(t)})}{\pi_{\omega_{\mathrm{old}}}(a^{(t)} \mid S^{(t)})}, \tag{14}$$

$$\mathcal{L}_{\mathrm{actor}} = -\mathbb{E}\left[\min\left(\rho_t(\omega)\hat{A}_t,\, \mathrm{clip}(\rho_t(\omega), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right] - \alpha\mathcal{H}(\pi_\omega), \tag{15}$$

$$\mathcal{L}_{\mathrm{critic}} = \mathrm{MSE}\big(V_\psi(S^{(t)}), R_t\big), \tag{16}$$

where $\hat{A}_t$ denotes the GAE advantage estimate, R_t is the return target, and $\mathcal{H}(\pi_\omega)$ is the entropy regularizer. The sampled point-wise actions form the intervention mask $A^{(t)}$, which is passed to the Bias-aware Refinement Module (BRM) for selective residual refinement.

A.4 BRM and BiasNet Architecture

The Bias-aware Refinement Module (BRM) performs constrained residual refinement on the points selected by AIA. Within BRM, BiasNet serves as the residual prediction component. Given the frozen prior feature F, the multi-scale prediction prior $P^{(t)}$, the uncertainty map $U^{(t)}$, and the intervention mask $A^{(t)}$, BiasNet first constructs a hybrid spatial-feature context to enrich the representation used for residual prediction. The intervention mask does not provide semantic information; instead, it serves as a gating signal that specifies where the predicted residual should be applied.

To provide both local geometric support and non-local semantic support, BiasNet constructs two complementary neighborhoods:

$$\mathcal{N}^s(i) = \mathrm{TopK}_{K_s}\left(-\|p_i - p_j\|_2\right), \tag{17}$$

$$\mathcal{N}^f(i) = \mathrm{TopK}_{K_f}\left(\frac{F_i^\top F_j}{\|F_i\|_2\|F_j\|_2}\right). \tag{18}$$

Here, $\mathcal{N}^s(i)$ is built from point coordinates and provides local geometric support, while $\mathcal{N}^f(i)$ is built from normalized frozen prior features and provides non-local semantic support. BiasNet aggregates neighbor information from these two neighborhoods and fuses the spatial and feature contexts using a lightweight attention-based fusion module. The fused context is denoted as:

$$H_i = \mathrm{AttnFuse}(H_i^s, H_i^f), \tag{19}$$

where H_i^s and H_i^f denote the aggregated spatial and feature contexts, respectively. This attention-based fusion adaptively combines local geometric consistency and semantic support, producing a richer contextual representation for residual prediction.

The enhanced context H is then combined with the multi-scale prediction prior $P^{(t)}$ and the uncertainty map $U^{(t)}$ for residual prediction. BiasNet uses a lightweight 1D convolutional residual head to predict the logit offset $\Delta Z^{(t)}$. Following the action-gated transition in Eq. (8), this residual is applied only to the agent-selected points indicated by $A^{(t)}$, while unselected points remain unchanged. This allows BiasNet to refine uncertain predictions without unnecessarily modifying reliable regions.

Refinement loss of BiasNet. The refinement loss guides BiasNet to correct uncertain predictions while preserving reliable semantic priors under sparse supervision. Below, we detail the computation and role of each loss term in the BiasNet objective.

- **Sparse semantic anchoring.** $\mathcal{L}_{\mathrm{CE}}$ is a cross-entropy loss that provides direct semantic supervision for residual refinement. It consists of supervised cross-entropy on the sparsely labeled points $\mathcal{I}_{\mathrm{sup}}$ and pseudo-label cross-entropy on the high-confidence unlabeled points $\mathcal{I}_{\mathrm{pseudo}}$. This term anchors the refined logits $Z^{(t+1)}$ to explicit semantic targets and prevents BiasNet from learning unconstrained corrections:

$$\mathcal{L}_{\mathrm{CE}} = \mathrm{CE}_{i\in\mathcal{I}_{\mathrm{sup}}}(Z_i^{(t+1)}, Y_i) + \lambda_{\mathrm{pseudo}}\mathrm{CE}_{i\in\mathcal{I}_{\mathrm{pseudo}}}(Z_i^{(t+1)}, \hat{Y}_i^{(t)}). \tag{20}$$

- **Distribution-level prior preservation.** $\mathcal{L}_{\mathrm{KL}}$ is a KL-divergence loss computed on high-confidence points $\mathcal{I}_{\mathrm{high}}$. It constrains the refined distribution $p_i^{(t+1)}$ to remain close to the pre-update distribution $p_i^{(t)}$, thereby preserving reliable semantic priors and discouraging unnecessary changes to confident regions:

$$\mathcal{L}_{\mathrm{KL}} = \mathbb{E}_{i\in\mathcal{I}_{\mathrm{high}}}\mathrm{KL}\left(p_i^{(t+1)} \,\|\, p_i^{(t)}\right). \tag{21}$$

- **Entropy-aware uncertainty reduction.** $\mathcal{L}_{\mathrm{ent}}$ is an entropy-aware loss computed on activated low-confidence points $\mathcal{I}_{\mathrm{low}}$. It encourages selected uncertain points to increase prediction confidence and reduce entropy after refinement, promoting effective correction on uncertain regions rather than passive preservation:

$$\mathcal{L}_{\mathrm{ent}} = -\mathbb{E}_{i\in\mathcal{I}_{\mathrm{low}}}\left[(c_i^{(t+1)} - c_i^{(t)}) + \eta\left(H(p_i^{(t)}) - H(p_i^{(t+1)})\right)\right]. \tag{22}$$

- **Neighborhood-guided regularization.** $\mathcal{L}_{\mathrm{neigh}}$ is a neighborhood-guided cross-entropy loss computed between the refined logits $Z^{(t+1)}$ and neighbor-induced targets Y_{neigh}. It encourages local prediction consistency and suppresses isolated noisy updates caused by unstable point-wise corrections:

$$\mathcal{L}_{\mathrm{neigh}} = \mathrm{CE}(Z^{(t+1)}, Y_{\mathrm{neigh}}). \tag{23}$$

Together, these four losses constrain BiasNet from complementary aspects: sparse semantic anchoring, reliable-prior preservation, uncertainty reduction, and local consistency regularization. This allows BiasNet to selectively refine uncertain predictions while preserving reliable semantic regions.

A.5 Detailed Reward Design

Overall reward. The reward is designed to guide the point-wise action selection of AIA. For point i at intervention step t, the reward consists of a label-driven feedback term and a confidence-aware shaping term:

$$r_i^{(t)} = r_{i,\mathrm{lab}}^{(t)} + r_{i,\mathrm{conf}}^{(t)}. \tag{24}$$

The label-driven term provides explicit correctness feedback on labeled points, while the confidence-aware term provides dense weak feedback on unlabeled points.

Label-driven feedback. For labeled points, the reward evaluates whether an intervention improves, degrades, or misses a prediction with available ground-truth supervision. Let $\hat{y}_i^{(t)}$ and $\hat{y}_i^{(t+1)}$ denote the predicted class before and after intervention, y_i denote the ground-truth label, and $a_i^{(t)} \in \{0, 1\}$ denote the action. As summarized in Table 8, the label-driven reward encourages activating points whose predictions can be improved, penalizes interventions that degrade already correct predictions, and discourages ignoring labeled errors. Thus, it provides sparse but reliable feedback for learning useful intervention actions.

Table 8: Definition of the label-driven reward.

Case	Before intervention	After intervention	Reward
Improve wrong prediction	$\hat{y}_i^{(t)} \neq y_i$	$\hat{y}_i^{(t+1)} = y_i, a_i^{(t)} = 1$	$+r_{\text{fix}}$
Degrade correct prediction	$\hat{y}_i^{(t)} = y_i$	$\hat{y}_i^{(t+1)} \neq y_i, a_i^{(t)} = 1$	$-r_{\text{break}}$
Ineffective intervention	$\hat{y}_i^{(t)} \neq y_i$	$\hat{y}_i^{(t+1)} \neq y_i, a_i^{(t)} = 1$	$-r_{\text{fail}}$
Miss labeled error	$\hat{y}_i^{(t)} \neq y_i$	$a_i^{(t)} = 0$	$-r_{\text{miss}}$
Otherwise	–	–	0

Confidence-aware shaping. Since point-level labels are extremely sparse, relying only on label-driven feedback provides insufficient signals for policy learning. We therefore introduce a confidence-aware shaping term for unlabeled points, using prediction confidence as a weak indicator of reliability. Let $c_i^{(t)} = \max_c p_{i,c}^{(t)}$ denote the prediction confidence, and let τ_{high} and τ_{low} denote the high-confidence and very-low-confidence thresholds. As summarized in Table 9, high-confidence predictions are encouraged to remain unchanged, while low-confidence predictions are encouraged to be activated for further refinement. The stronger penalty on missed very-low-confidence points further prevents the policy from collapsing into a low-activation strategy under sparse label feedback.

Table 9: Definition of the confidence-aware shaping reward.

Case	Condition	Reward
Activate high-confidence point	$c_i^{(t)} > \tau_{\text{high}},\ a_i^{(t)} = 1$	$-r_{\text{high}}$
Preserve high-confidence point	$c_i^{(t)} > \tau_{\text{high}},\ a_i^{(t)} = 0$	$+r_{\text{stable}}$
Activate low-confidence point	$c_i^{(t)} \leq \tau_{\text{high}},\ a_i^{(t)} = 1$	$+r_{\text{low}}$
Miss uncertain point	$\tau_{\text{low}} \leq c_i^{(t)} \leq \tau_{\text{high}},\ a_i^{(t)} = 0$	$-r_{\text{mid}}$
Miss very-low-confidence point	$c_i^{(t)} < \tau_{\text{low}},\ a_i^{(t)} = 0$	$-r_{\text{verylow}}$

B Implementation Details

We implement PARI with OpenPoints using AADNet as the default backbone. Training follows a two-stage procedure: the base segmentation network is first trained with sparse annotations to provide an initial semantic prior, and is then loaded from the pretrained checkpoint and frozen during Stage-II intervention training, where only AIA, BiasNet, and auxiliary modules are optimized. AIA uses lightweight two-layer Actor/Critic MLPs with hidden dimension 32 and a binary action space for activation and non-activation. We optimize the policy with PPO using a rollout length of $T = 2$, sample actions from a Bernoulli/categorical distribution, and evaluate single-step intervention as an efficient inference option. The contextual prior is extracted with the pyramid pooling variant of PPM. All methods adopt the same sparse annotation protocol, preprocessing, and sampling strategy, and are implemented in PyTorch; S3DIS and ScanNetV2 experiments are conducted on a single NVIDIA A100 GPU, while SemanticKITTI experiments are conducted on a single NVIDIA Tesla H20 GPU.

C Code and Datasets

C.1 Code

We will release the code and experimental configurations after publication to facilitate reproducibility.

C.2 Datasets and Annotation Protocol Details

Sparse annotation protocol. For all weakly supervised experiments, we construct sparse point-level annotations from the original dense labels of each dataset. Only the selected sparse labels are used during training, while all remaining points are treated as unlabeled. Dense labels are strictly used only for evaluation.

For the uniform annotation setting, S3DIS is evaluated under three sparse annotation ratios, including 1%, 0.1%, and 0.01%. For ScanNetV2, we evaluate both the 1% sparse annotation setting and the extreme 20pts setting, where only 20 labeled points are retained per scene. For the non-uniform annotation setting on S3DIS, labeled points are sampled with biased class or region distributions to simulate imbalanced annotation, following the ratios used in the main experiments.

Label-aware sampling. Since point cloud networks usually perform point sampling or downsampling during training, sparse labels may be accidentally removed, especially under extremely low annotation ratios. To avoid changing the effective annotation budget, we adopt label-aware sampling during training: annotated points are preserved in the sampled subset whenever possible, and the remaining positions are filled with unlabeled points using the standard sampling strategy.

Fair comparison. The same annotation generation, preprocessing pipeline, and label-aware sampling strategy are applied to all compared methods. This protocol does not introduce additional supervision, but ensures that each method is trained with the same sparse annotation budget and comparable labeled-point availability.

D Additional Experiments and Analysis

D.1 Ablation Study for Top-k Neighborhood Size

This experiment evaluates the effect of explicit neighborhood modeling in BiasNet. We compare the variant without neighborhood aggregation with different top-k settings on S3DIS under 0.1% annotations. As shown in Table 10, introducing neighborhood information consistently improves residual refinement over the no-neighborhood variant, indicating that local context provides useful structural cues for estimating reliable residual updates.

Table 10: Ablation study for the top-k neighborhood size in BiasNet on S3DIS under 0.1% annotations.

Setting	w/o neigh.	$K = 8$	$K = 16$	$K = 24$
mIoU	67.2	67.3	**67.9**	67.5

The best performance is achieved with $K = 16$, which reaches 67.9 mIoU and improves the no-neighborhood variant by 0.7 mIoU. Using a smaller neighborhood, such as $K = 8$, may provide insufficient local structural information, while a larger neighborhood, such as $K = 24$, may introduce semantically inconsistent points, especially around object boundaries. We therefore use $K = 16$ as the default setting, as it provides a better balance between local context coverage and semantic relevance.

D.2 Ablation Study for the Number of Intervention Steps

We examine the choice of intervention steps during training and inference on S3DIS. Table 11 compares configurations with one, two, and three training steps under 0.1% annotations. The two-step configuration achieves the highest mIoU of 67.9, compared with 67.4 and 67.5 for the one-step and three-step configurations, respectively. These results inform our choice of a two-step training configuration.

Table 11: Comparison of training configurations on S3DIS with 0.1% annotations. Each model is evaluated using the same number of intervention steps as during training.

Training steps T_{train}	1	2	3
mIoU	67.4	**67.9**	67.5

We further evaluate each model trained with $T_{\text{train}} = 2$ using one or two inference steps. As shown in Table 12, the second step produces only marginal changes under 1% and 0.1% annotations and reduces performance under 0.01% annotations. These results support single-step inference as a lower-cost alternative, since the second step provides no consistent performance benefit.

Table 12: Effect of inference steps on S3DIS. Training is fixed at $T_{\text{train}} = 2$.

Annotation	$T_{\text{test}} = 1$	$T_{\text{test}} = 2$
1%	69.403	69.402
0.1%	67.842	67.861
0.01%	61.572	61.423

D.3 Ablation Study for Reward Coefficients

This experiment evaluates whether PARI is sensitive to the numerical choices of reward coefficients. We vary the coefficients of the label-driven reward, high-confidence preservation reward, and low-confidence exploration reward on S3DIS under 0.1% annotations.

Table 13: Sensitivity analysis of reward coefficients on S3DIS under 0.1% annotations.

Group	Configuration	mIoU
Label-driven	$(7.0, -4.0, -0.5, -0.5)$ for $(r_{\text{fix}}, -r_{\text{break}}, -r_{\text{fail}}, -r_{\text{miss}})$	67.4
Label-driven	$(4.0, -7.0, -0.5, -0.5)$	67.8
Label-driven	$(5.0, -5.0, -0.5, -0.5)$	**67.9**
Label-driven	$(5.0, -5.0, -0.5, -1.0)$	67.3
High-conf.	$(-2.5, 0.05)$ for $(-r_{\text{high}}, r_{\text{stable}})$	67.4
High-conf.	$(-0.8, 0.3)$	67.4
High-conf.	$(-0.5, 0.5)$	67.5
Low-conf.	$(-1.2, -0.5, 0.5)$ for $(-r_{\text{verylow}}, -r_{\text{mid}}, r_{\text{low}})$	67.6
Low-conf.	$(-0.4, -0.15, 0.15)$	67.6

As shown in Table 13, most coefficient configurations achieve stable performance between 67.3 and 67.8 mIoU, while the default balanced setting obtains the best result of 67.9 mIoU. This indicates that PARI is not highly sensitive to exact reward magnitudes, but benefits from the complementary structure of different reward terms.

For the label-driven reward, the best result is obtained when successful correction and destructive intervention are balanced. Over-emphasizing correction rewards may encourage aggressive updates, while an overly strong missing penalty may push the policy to intervene too broadly. For the high-confidence reward, different settings lead to similar performance, suggesting that its main role is to preserve reliable predictions rather than dominate policy learning. For the low-confidence reward, the stable results show that encouraging intervention on uncertain regions provides useful dense feedback under sparse supervision. Overall, the reward design works by combining sparse supervised correctness feedback, high-confidence preservation, and low-confidence exploration. The default configuration provides the best balance among these signals and is used in all main experiments.

In the evaluated runs, we observed no NaN/Inf values or loss explosion, and the activation rate converged to a finite, nonzero value. Critic loss decreased from 1.212 to 0.103 under 0.01% annotations and from 0.664 to 0.197 under 1% annotations.

D.4 Confidence–Reliability Correlation on Unlabeled Predictions

This analysis examines whether prediction confidence can serve as an effective reliability indicator for unlabeled points, thereby supporting our confidence-aware reward design. We group unlabeled points according to their Stage-I prediction confidence and compute the prediction correctness within each confidence interval. Ground-truth labels are used only for this post-hoc analysis and are never accessed during training.

As shown in Fig. 5, prediction correctness increases consistently with confidence. This trend supports our confidence-aware reward design: high-confidence regions are treated as preservation targets to avoid unnecessary modifications, while low-confidence regions provide activation signals for policy intervention. Thus, PARI uses the confidence structure of unlabeled predictions as weak feedback for policy learning, rather than directly treating unlabeled predictions as dense supervision. However, high-confidence suppression is a soft penalty and does not guarantee correction of high-confidence

errors. On S3DIS with 1% annotations, all 11,418 beam points above the suppression threshold were misclassified by Stage I, and none was corrected by PARI; the beam IoU remained 0.0. This illustrates that confidence-based suppression may preserve erroneous predictions when the underlying semantic evidence is insufficient.

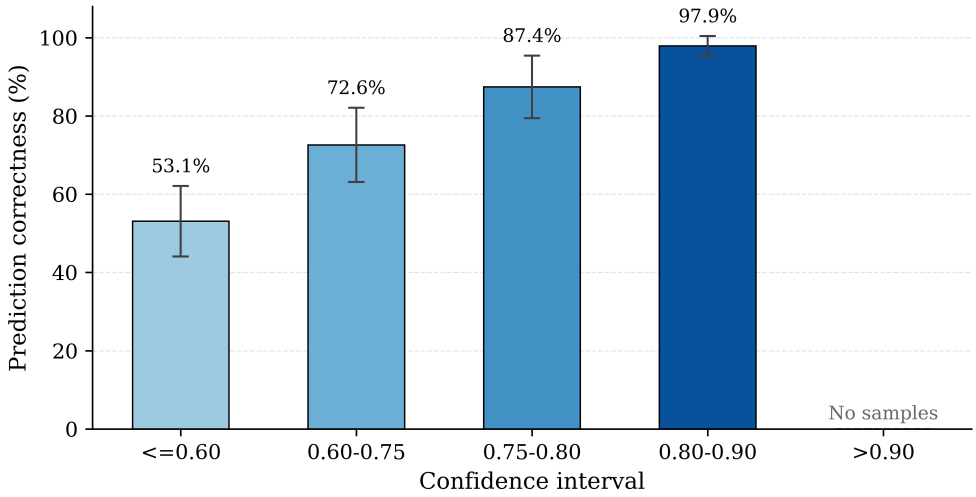

Figure 5: Reliability of unlabeled predictions across different confidence intervals. We report the average Stage-I prediction correctness within each confidence interval over 68 rooms. Prediction correctness increases consistently with confidence, indicating that confidence is strongly correlated with prediction reliability.

D.5 Reward-guided Actor Focus Analysis

This analysis examines how the learned AIA allocates interventions across unlabeled points with different confidence levels. We group Stage-I unlabeled predictions by confidence and report three statistics for each interval: *Stage-I weak-region density*, which measures the proportion of potentially unreliable predictions; *Actor focus rate*, which measures the proportion of points selected for intervention by the AIA; and *Aligned weak-region focus*, which measures the overlap between selected points and weak regions. This allows us to observe whether the Actor tends to select unreliable regions rather than distributing interventions uniformly over all unlabeled points.

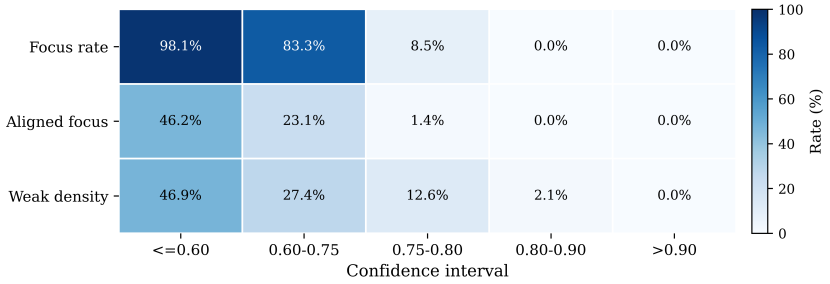

Figure 6: Reward-guided actor focus analysis on unlabeled points. AIA allocates more interventions to low-confidence regions and reduces actions on high-confidence predictions.

As shown in Fig. 6, the AIA selects more points from low-confidence intervals, where Stage-I predictions are more likely to be unreliable. When confidence is below 0.60, the Actor focus rate reaches 98.1%, and it remains high at 83.3% in the 0.60–0.75 interval. As confidence increases, the focus rate drops quickly and becomes nearly zero when confidence exceeds 0.80. This indicates that the Actor learns to focus on potentially unreliable regions while avoiding frequent interventions on high-confidence predictions.

D.6 Additional Results on SemanticKITTI

We further evaluate PARI on SemanticKITTI to examine its transferability from indoor point clouds to outdoor LiDAR scenes. Following the same weak-supervision protocol, only sparse point-level labels are used for training, while the remaining points are treated as unlabeled. PARI is applied on top of the Stage-I base model under the same preprocessing and sampling strategy.

Table 14: Semantic segmentation results on SemanticKITTI under sparse annotations.

Method	Annotation	mIoU	Gain
Base model	0.1%	47.84	–
Base model + PARI	0.1%	48.05	0.21

As shown in Table 14, PARI improves the base model from 47.84 to 48.05 mIoU under 0.1% annotations. Although the gain is moderate, it suggests that the proposed intervention mechanism can also provide benefits on outdoor LiDAR scenes.

D.7 Effect of High-confidence Pseudo-labels

We analyze the quality of pseudo-labels and observe that high-confidence predictions achieve an accuracy exceeding **98.5**% across all weakly supervised settings. Here, the statistics are computed on high-confidence pseudo-labels selected from the training set for use in BRM. This result demonstrates that the reliability estimation strategy adopted in the Bias-aware Refinement Module (BRM) is well-founded. To further verify that simply leveraging high-confidence points is insufficient to improve segmentation performance, we use the same set of high-confidence pseudo-labels employed in BRM to train the base model. The results show comparable performance (**68.27**% vs. **68.30**%), indicating that solely relying on high-confidence points does not lead to meaningful performance gains.

D.8 Results across random seeds

We evaluate PARI across multiple random seeds and report the mean and sample standard deviation of mIoU in Table 15. We perform four runs on S3DIS and three runs each on ScanNetV2 and SemanticKITTI.

Table 15: Results across random seeds. std denotes the sample standard deviation.

Dataset	Annotation	Runs	mIoU (mean $\pm$ std)
S3DIS	1%	4	69.39 ± 0.08
ScanNetV2	1%	3	67.15 ± 0.07
SemanticKITTI	0.1%	3	48.04 ± 0.03

D.9 Comparison with reproduced DGNet

We reproduce DGNet [22] using the PointNeXt [26] backbone, S3DIS Area 5, 0.04 m voxelization, and uniformly sparse annotations. As shown in Table 16, PARI exceeds DGNet by 1.69 mIoU at 1% annotations, matches its result at the reported precision at 0.1%, and falls behind by 0.49 mIoU at 0.01%. Thus, neither method consistently outperforms the other across all evaluated annotation ratios.

Table 16: Comparison with our reproduced DGNet on S3DIS Area 5 under uniformly sparse annotations. Results are reported in mIoU, where * denotes results reproduced by us.

Method	1%	0.1%	0.01%
DGNet*	67.71	67.86	61.91
PARI	69.40	67.86	61.42

The two methods are potentially complementary rather than mutually exclusive. DGNet guides Stage-I representation learning through class distributions and prototypes, while PARI learns where to intervene and how to refine residual errors given frozen Stage-I priors. A base model could first be trained with DGNet-style objectives, after which its frozen features and logits could be supplied to PARI. Since AIA does not depend on a specific Stage-I training objective, this combination would not require changing PARI's core formulation. However, we have not evaluated DGNet+PARI and therefore identify this combination as future work.

D.10 Additional Qualitative Results

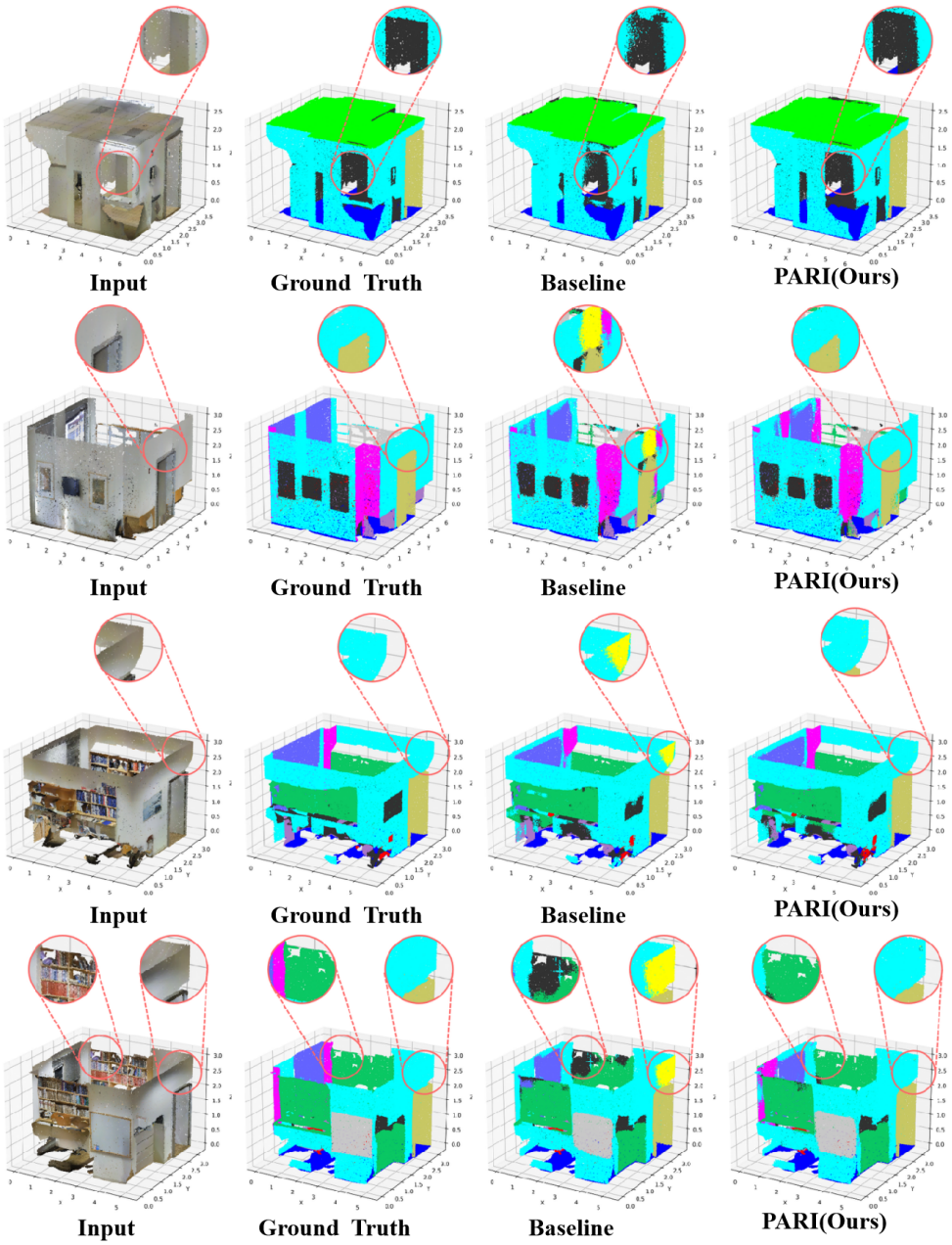

Figure 7: Additional qualitative results on challenging indoor scenes. PARI improves predictions in semantic boundaries and structurally complex regions, indicating that the intervention mechanism helps refine difficult point cloud regions while preserving reliable areas.

E Theoretical Analysis and Discussion

E.1 How Confidence-aware Reward Shaping from Unlabeled Regions Affects Agent Decisions

We analyze how confidence-aware reward shaping on unlabeled regions affects the decision behavior of AIA. In PPO, the policy update is determined by the advantage estimate. When only labeled feedback is used, the step reward is

$$r_t^{\text{sup}} = \frac{1}{N}\sum_{i=1}^{N} r_{i,\text{lab}}^{(t)}. \tag{25}$$

After introducing confidence-aware reward shaping, the step reward becomes

$$r_t^{\text{shape}} = \frac{1}{N}\sum_{i=1}^{N}\left(r_{i,\text{lab}}^{(t)} + r_{i,\text{conf}}^{(t)}\right), \tag{26}$$

where $r_{i,\text{conf}}^{(t)}$ is triggered by the confidence-based reward rule on unlabeled points.

The temporal-difference (TD) residual and generalized advantage estimation (GAE) advantage are computed as

$$\delta_t = r_t + \gamma V_\psi(S^{(t+1)}) - V_\psi(S^{(t)}), \tag{27}$$

$$\hat{A}_t = \sum_{l=0}^{T-t-1} (\gamma\lambda)^l \delta_{t+l}. \tag{28}$$

Thus, under the same value estimates, the advantage change caused by confidence-aware reward shaping is

$$\hat{A}_t^{\text{shape}} - \hat{A}_t^{\text{sup}} = \sum_{l=0}^{T-t-1} (\gamma\lambda)^l \frac{1}{N}\sum_{i=1}^{N} r_{i,\text{conf}}^{(t+l)}. \tag{29}$$

This shows that confidence-aware rewards from unlabeled regions directly change the advantage signal used for PPO policy updates. A positive confidence reward increases the advantage of the corresponding action and encourages similar decisions in future states, while a negative confidence reward decreases the advantage and suppresses similar decisions. Therefore, low-confidence exploration rewards guide AIA to select uncertain regions with higher correction potential, whereas high-confidence preservation rewards discourage unnecessary interventions on reliable predictions. In this way, confidence-aware reward shaping converts unlabeled regions into policy-level feedback and directly affects where the agent decides to intervene.

E.2 How Policy-based Point Selection Improves Intervention Utility

We analyze why policy-based point selection can be more effective than heuristic selection strategies. In PARI, the utility of selecting a point is not determined only by its current uncertainty, because an intervention changes the logits, the next prediction state, and the subsequent rewards. Therefore, point selection should be viewed as maximizing long-term intervention utility rather than performing one-step ranking.

For point i at state $S^{(t)}$, we define its long-term intervention utility as the value difference between activating and not activating this point:

$$g_i^{(t)} = Q^\pi(S^{(t)}, A_i^{(t)} = 1) - Q^\pi(S^{(t)}, A_i^{(t)} = 0), \tag{30}$$

where

$$Q^\pi(S^{(t)}, A^{(t)}) = \mathbb{E}_\pi\left[\sum_{l=t}^{T-1} \gamma^{l-t} r(S^{(l)}, A^{(l)}) \mid S^{(t)}, A^{(t)}\right]. \tag{31}$$

A point with $g_i^{(t)} > 0$ has positive intervention utility, while a point with $g_i^{(t)} < 0$ is unhelpful or harmful to intervene.

Assume that selected positive-utility points have expected gain $\mu_+ > 0$, and selected non-positive-utility points have expected loss $\mu_- > 0$, with these conditional means shared across the compared policies. If a selection policy chooses a subset in which the proportion of positive-utility points is p, then its expected utility satisfies:

$$\mathbb{E}[g] = p\mu_+ - (1-p)\mu_-. \tag{32}$$

Therefore, for two selection policies π_1 and π_2 with positive-utility proportions p_1 and p_2, we have:

$$\mathbb{E}[g_{\pi_1}] - \mathbb{E}[g_{\pi_2}] = (p_1 - p_2)(\mu_+ + \mu_-). \tag{33}$$

This result shows that the expected utility of point selection increases with the proportion of selected positive-utility points. Random selection cannot explicitly distinguish useful and harmful interventions, all-point intervention may include many reliable points with negative utility, and entropy ranking only considers current uncertainty without modeling future state changes. In contrast, AIA is optimized through reward feedback and PPO advantage estimates, which reinforce actions leading to higher long-term utility and suppress ineffective or harmful actions. Thus, the potential benefit of policy-based selection comes from increasing the probability of selecting points with positive long-term intervention utility, rather than simply selecting more points.

E.3 How Selective Intervention Reduces Expected Prediction Error

We analyze when selective residual intervention can reduce prediction error in expectation. Let the point-wise error change after intervention be

$$\Delta e_i = e_i^{(t+1)} - e_i^{(t)}. \tag{34}$$

A negative Δe_i indicates improvement, while a positive value indicates degradation. Assume that residual refinement decreases the error of high-risk points in expectation, but may increase the error of reliable points due to unnecessary modification:

$$\mathbb{E}[\Delta e_i \mid i \in \mathcal{H}] \le -\alpha, \qquad \mathbb{E}[\Delta e_i \mid i \in \mathcal{R}] \le \beta, \qquad \alpha, \beta > 0. \tag{35}$$

Here, $\mathcal{H}$ and $\mathcal{R}$ denote high-risk and reliable points within the selected subset, respectively.

Suppose AIA selects M points for intervention, and the proportion of high-risk points among them is p. Then the expected error change of selective intervention is bounded by

$$\mathbb{E}[\Delta E_{\text{sel}}] \le -pM\alpha + (1-p)M\beta. \tag{36}$$

Therefore, selective intervention reduces the expected error when

$$p > \frac{\beta}{\alpha + \beta} \quad \Rightarrow \quad \mathbb{E}[\Delta E_{\text{sel}}] < 0. \tag{37}$$

This result shows that selective intervention reduces the expected prediction error when the selected subset contains a sufficiently high proportion of truly high-risk points. This condition is supported by the reward-guided Actor focus analysis in Sec. D.5, where AIA allocates most interventions to low-confidence intervals with higher weak-region density and suppresses actions on high-confidence predictions. For example, the Actor focus rate reaches 98.1% when confidence is below 0.60 and remains 83.3% in the 0.60–0.75 interval, but becomes nearly zero when confidence exceeds 0.80. This indicates that the learned policy concentrates its intervention budget on potentially unreliable regions rather than selecting points uniformly. Therefore, AIA increases the proportion of high-risk points in the selected subset, making the condition $p > \beta/(\alpha + \beta)$ more likely to hold and allowing selective residual intervention to reduce prediction error in expectation.

E.4 Discussion: From Feedback-driven Post-training to Semantic Intervention

PARI is conceptually related to feedback-driven post-training in large models. Both start from an existing model prior and use feedback signals to further improve model behavior, rather than relying solely on supervised fitting. In this sense, PARI can be viewed as a feedback-driven refinement process built upon a learned semantic prior.

However, the target and structure of PARI differ from standard response-level post-training. Large-model post-training typically optimizes complete responses or preference pairs, whereas PARI

operates on dense 3D predictions, where different points have different reliability and refinement value. Therefore, the key problem is not only how to improve the overall prediction, but also where feedback-guided refinement should be applied.

PARI further formulates this refinement as a multi-step semantic intervention process. Each intervention updates the current prediction state and changes subsequent uncertainty, multi-scale prediction priors, and policy decisions. AIA determines where refinement should be activated, while BiasNet performs residual refinement on the selected points. Thus, PARI extends the idea of feedback-driven optimization from response-level improvement to point-level, multi-step semantic refinement for weakly supervised 3D segmentation.

F Limitations and Future Work

Despite the consistent improvements achieved by the proposed PARI across various weakly supervised settings, several limitations remain. First, the current framework relies on a two-stage training pipeline and a multi-step intervention process. While this design stabilizes policy learning, it also introduces additional training complexity and hyperparameter tuning overhead. Second, the decision quality of the RL agent is still influenced by the state representation and reward design; when facing more complex scene structures or long-tailed category distributions, there remains room for further improvement in policy learning. In addition, the current method has been systematically validated primarily on indoor point cloud semantic segmentation benchmarks. Its generalization ability to larger-scale outdoor scenarios, cross-sensor data, and more open annotation settings still requires further investigation.

Future work can be explored along several directions. First, more efficient policy learning mechanisms could be developed to reduce the overhead introduced by multi-step interventions, thereby improving both training and inference efficiency. Second, richer state representations and reward formulations can be designed to better adapt intervention strategies to complex topologies, long-tailed distributions, and open environments. Third, recent progress in reinforcement learning for multimodal learning and reasoning [45, 46, 35, 38, 51] motivates extending PARI to multimodal 3D understanding. We will explore incorporating image and language cues into PARI's policy state to learn when and where cross-modal evidence should guide residual refinement. Extensions to instance segmentation, panoptic segmentation, and 3D scene completion also warrant investigation.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The motivation for proposing the PARI framework, the specific methods, and the experimental conclusions on S3DIS and ScanNetV2 are clearly stated in the abstract and introduction, which accurately reflect the actual content of the paper.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We objectively discuss the limitations of our work in the Appendix section.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: The paper provides analytical justifications in the appendix to support the proposed design. These analyses cover how confidence-aware reward shaping affects PPO advantage estimation, why policy-based point selection improves long-term intervention utility, and under what condition selective residual intervention reduces expected prediction error. The relevant assumptions and derivations are stated together with each analysis.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: We provide detailed descriptions of the implementation details, ensuring the full reproducibility of our experiments.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in

some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [No]

   Justification: We will provide the complete open-source code, training scripts, and pre-trained model weights upon acceptance.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Core training hyperparameters, network architecture details, reward thresholds for the PPO algorithm, and the weights for the multi-objective loss are detailed in our manuscript.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Table 15 in Appendix D.8 reports the mean and sample standard deviation of mIoU across random seeds: four runs on S3DIS with 1% annotations, three runs on ScanNetV2 with 1% annotations, and three runs on SemanticKITTI with 0.1% annotations.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.

- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: All experiments are conducted in PyTorch. S3DIS and ScanNetV2 experiments are performed on a single NVIDIA A100 GPU, while SemanticKITTI experiments are conducted on a single NVIDIA Tesla H20 GPU.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research in this paper strictly adheres to the NeurIPS Code of Ethics. The datasets used in our experiments are fully public academic datasets, and there are no privacy violations or improper usage.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: The weakly supervised 3D point cloud segmentation technology explored in this paper can significantly reduce data annotation costs in robotic perception and autonomous driving, bringing primarily positive industrial and societal impacts without obvious negative consequences.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The methods and models proposed in this paper are primarily for fundamental academic research in 3D indoor scene understanding, posing no high risks or potential for being misused to generate fake or dangerous content.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: The S3DIS and ScanNetV2 datasets, as well as the OpenPoints codebase used in our work, are open-source academic assets. Appropriate citations and acknowledgments are provided in the references.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [N/A]

Justification: This research proposes a new algorithmic framework and does not collect, annotate, or release any new dataset resources or other new assets at submission time.

Guidelines:

- The answer [N/A] means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [N/A]

Justification: This study evaluates algorithms entirely based on existing open-source academic datasets and does not involve any crowdsourcing experiments or human subjects research.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: Since this research does not involve any human subjects or personal privacy data, approval from an Institutional Review Board (IRB) is not required.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: Large language models were only used to assist in improving English writing and polishing grammar.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.